Patent

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

COPY OF PAPERS ORIGINALLY FILED

As a below-named inventor, I hereby declare that:

My residence, post office address, and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

A SCALABLE SYSTEM FOR CLUSTERING OF LARGE DATABASES MIXED DATA ATTRIBUTES

the specification of which:

    _____ is attached hereto.

    _____ was filed on as United States Patent Application Number and was amended on ___ (if applicable).

    ___X___ was filed on March 29, 1999 as PCT International Application Number ___PCT/US99/06717___.

I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose to the United States Patent and Trademark Office information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56. Under that section, information is material to patentability when it is not cumulative to information already of record or being made of record in the application, and (1) it establishes, by itself or in combination with other information, a prima facie case of unpatentability of a claim; or (2) it refutes, or is inconsistent with, a position the applicant takes in: (i) opposing an argument of unpatentability relied on by the United States Patent and Trademark Office, or (ii) asserting an argument of patentability.

I hereby claim foreign priority benefits under Title 35, United States Code, § 119(a)-(d) or § 365(b) of any foreign application(s) for patent or inventor's certificate, or § 365(a) of any PCT international application(s) which designated at least one country other than the United States, listed below. I have also identified below any corresponding foreign application(s) for patent or inventor's certificate or any corresponding PCT international application(s) having a filing date before that of the application(s) on which priority is claimed.

Prior Foreign Application(s)

Number	Country	Filing Date (mm/dd/yyyy)	Priority Claimed (Yes or No)
PCT/US99/06717	PCT	29 March 1999	Yes

I hereby claim the benefit under Title 35, United States Code, § 119(e) of any United States provisional application(s) listed below.

United States Provisional Application(s)

Application Number	Filing Date (mm/dd/yyyy)
60/086,410	May 22, 1998

I hereby claim the benefit under Title 35, United States Code, § 120 of any United States application(s) or § 365(c) of any PCT international application(s) designating the United States, listed below. Insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States or PCT international application(s) in the manner provided by the first paragraph of Title 35, United States Code, § 112, I acknowledge the duty to disclose to the United States Patent and Trademark Office information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, § 1.56 which became available between the filing date of the prior application(s) and the national or PCT international filing date of this application.

United States Application(s) Upon Which Priority is Claimed

Application Number	Filing Date (mm/dd/yyyy)	Status (abandoned, pending, patent number)
09/083,906	May 22, 1998	Pending

I hereby appoint the following attorney(s) and/or agent(s), all of whom are associated with Watts, Hoffmann, Fisher & Heinke Co., L.P.A., having offices at 1100 Superior Avenue East, Suite 1750, Cleveland, Ohio 44114-2518, to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith:

Thomas E. Fisher, Reg. No. 18,271
Stephen J. Schultz, Reg. No. 29,108
Lowell L. Heinke, Reg. No. 19,471
John R. Hlavka, Reg. No. 29,076
George L. Pinchak, Jr., Reg. No. 37,697

Linn J. Raney, Reg. No. 23,078
Paul A. Serbinowski, Reg. No. 34,429
Richard A. Sharpe, Reg. No. 34,722
James G. Watterson, Reg. No. 20,180
Peter R. Hagerty, Reg. No. 42,618
Kenneth J. Smith, Reg. No. 45,115

and the following attorney(s), all of whom are associated with Microsoft Corporation having an address of One Microsoft Way, Redmond, Washington 98052-6399;

Daniel D. Crouse, Reg. No. 32,022   Katie E. Sako, Reg. No. 32,628

to prosecute this application and to transact all business in the United States Patent and Trademark Office connected therewith, and I hereby designate the foregoing Stephen J. Schultz, Reg. No. 29,108, principal attorney.

This appointment shall include all power to prosecute and transact all business relating to all applications corresponding to this application in all countries, including all regional and international patent offices, such as, but not limited to, the European Patent Office and all offices and bureaus established in accordance with the Patent Cooperation Treaty.

I hereby further designate and appoint any officer of Watts, Hoffmann, Fisher & Heinke Co., L.P.A. my attorney in fact with full power of substitution and revocation, including power to designate a substitute principal attorney.

Address all correspondence to:

WATTS, HOFFMANN, FISHER & HEINKE CO., L.P.A.
1100 Superior Ave., Ste. 1750
Cleveland, Ohio 44114-2518

Direct all telephone calls to Stephen J. Schultz at (216) 241-6700 and direct all facsimile correspondence to (216) 241-8151.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under § 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of this application or any patent issued thereon.

Full Name of Sole or First Joint Inventor: Usama Fayyad	
Inventor's Signature:	Date:
Residence: Mercer Island, Washington	Citizenship: United States of America
Post Office Address: 9705 SE 43rd Street, Mercer Island, WA 98040	

Full Name of Second Joint Inventor: Paul S. Bradley	
Inventor's Signature:	Date:
Residence: Seattle, Washington	Citizenship: United States of America
Post Office Address: 600 Lake Washington Blvd., East, Apt #C, Seattle, WA 98112	

Full Name of Third Joint Inventor: Cory A. Reina	
Inventor's Signature:	Date:
Residence: Kirkland, Washington	Citizenship: United States of America
Post Office Address: 12107 86th Avenue, NE, Kirkland, WA 98034	

A Scalable System for Clustering of Large Databases Having Mixed Data Attributes

Cross Reference to Related Applications

The present invention is a continuation in part of co-pending United States patent application serial number 09/083,906 entitled "A Scalable System for Expectation Maximization Clustering of Large Databases" to Fayyad et al having a filing date of May 22, 1998, and which is assigned to the assignee of the present application and is incorporated herein by reference.

Priority is claimed from provisional United States patent application serial
10   number 60/086,410 filed May 22, 1998.

Field of the Invention

The present invention concerns database analysis and more particularly concerns an apparatus and method for clustering of data into groups that capture
15   important regularities and characteristics of the data.

Background Art

Large data sets are now commonly used in most business organizations. In fact, so much data has been gathered that asking even a simple question about the data
20   has become a challenge. The modern information revolution is creating huge data stores which, instead of offering increased productivity and new opportunities, can overwhelm the users with a flood of information. Tapping into large databases for even simple browsing can result in a return of irrelevant and unimportant facts. Even people who do not 'own' large databases face the overload problem when accessing
25   databases on the Internet. A large challenge now facing the database community is how to sift through these databases to find useful information.

Existing database management systems (DBMS) perform the steps of reliably storing data and retrieving the data using a data access language, typically SQL. One major use of database technology is to help individuals and organizations make
30   decisions and generate reports based on the data contained in the database.

An important class of problems in the areas of *decision support* and *reporting* are clustering (segmentation) problems where one is interested in finding groupings (clusters) in the data. Data clustering has been used in statistics, pattern recognition,